

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 1, 2003, 08:49:06 ; Search time 93.5 Seconds
(without alignments)
6121.913 Million cell updates/sec

Title: US-09-768-781-2

Perfect score: 2543

Sequence: 1 atgaacacaagaccacaaca.....caaggcaagtggtgtctga 1389

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DRV=xlp
-Q/cgn2_1/USFTO.spool/US09768781/runat_01042003_084730_4939/app_query.fasta_1.1543
-DB=SPTRMBL_21 -QFWT=fastan -SURFIX=n2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09768781 @CN_1_1_158 @runat_01042003_084730_4939 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTRMBL_21:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	953.5	37.5	446	11	Q9QXY7 mus musculus

ALIGNMENTS

RESULT 1

Q9QXY7 ID Q9QXY7 PRELIMINARY; PRT; 446 AA.

AC Q9QXY7; DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE KX antigen (1810038K39RIK protein).
GN XXH OR XK OR 1810038K19RIK.
OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=SKELETAL MUSCLE;
RX MEDLINE=20009522; PubMed=10541802;

RA Collec E., Colin Y., Carbonnet F., Hattab C., Bertrand O.,

RA Carton J.P., Kim C.L.;

RT "Structure and expression of the mouse homologue of the XK gene."
RL Immunogenetics 50:16-21(1999).

SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=SKELETAL MUSCLE;
RC Le Van Kim C., Collec E., Colin Y.;

Q9h6d3 homo sapien
Q96p28 homo sapien
Q8tba0 homo sapien
O17386 caenorhabdi
Q8r1i8 mus musculus
Q9sak8 arabidopsis
Q9v2c2 pyrococcus
Q85329 oryza sativ
Q942c7 oryza sativ
Q9b5z8 schistosoma
Q9nug5 homo sapien
Q94ux1 schistosoma
Q94ux2 schistosoma
Q94ux6 schistosoma
Q9kcw3 bacillus ha
Q9c8a6 arabidopsis
Q9p6n5 schistosacch
Q94ux8 schistosoma
O47543 chlamydomon
Q98qc2 mycoplasma
O21706 pygathrix b
Q9b509 tetrodonop
O06039 lactococcus
Q9sa37 arabidopsis
O21845 pygathrix b
O21707 pygathrix b
Q955t7 cheirogaleu
Q8rv12 arabidopsis
O21708 pygathrix b
Q9qj11 cercopithec
Q8xm39 clostridium
Q8zqn3 salmonella
Q955t8 cheirogaleu
Q97e21 clostridium
Q19463 caenorhabdi
Q994e7 porcine ade
Q9p6w3 neurospora
Q8x7v5 escherichia
Q99xh1 streptococc
Q8z874 salmonella
Q963l7 schistosoma
Q955t9 cheirogaleu
Q9xdk3 bacteroides
Q8sey6 bombyx mand

Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.

```

RL  [3]
RN  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX  MEDLINE=21085660; PubMed=11217851;
RA  Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA  Arakawa T., Hata A., Fukuishi Y., Konno H., Adachi J., Fukuda S.,
RA  Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakura I.,
RA  Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA  Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA  Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA  Kuehl P., Lewis S., Matsuo Y., Nikaado I., Pesole G., Quackenbush J.,
RA  Schraml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA  Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA  Blake J., Boffelli D., Bojunga N., Carrinci P., de Bonaldo M.F.,
RA  Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA  Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA  Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA  Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA  Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA  Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA  Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA  Hayaishizaki Y.;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
DR  ENBL; AF155511; AAF14527.1; -.
DR  ENBL; AK007734; BAB25222.1; -.
DR  MGD; MGI:103569; Xkh.
SQ  SEQUENCE 446 AA; 51114 MW; D785FB7B9E28B98B CRC64;

```

Alignment Scores:

```

Pred. No.: 5,39e-83 Length: 446
Score: 953.50 Matches: 181
Percent Similarity: 64.72% Conservative: 85
Best Local Similarity: 44.04% Mismatches: 134
Query Match: 37.50% Indels: 11
DB: 11 Gaps: 4

```

US-09-768-781-2 (1-1389) x Q9QXV7 (1-446)

```

QY 136 TTTCATTTAGCATCTCTTCCACCTTTTGTACTGTGGGAGGCTGTCATCTGCTTG 195
DB 3 PheProAlaSerValIleAlaSerValPheLeuPheValAlaGluThrAlaAlaLeu 22
QY 196 TACATGTTAGATCTATCGAAGAATAGTGAACCTTACTGATGACATACACCTTTCT 255
DB 23 TyrLeuSerSerThrTyrArgSerAlaGlyAspArgMetTrpGlnValLeuLeu 42
QY 256 TTCTTTATGTTTTCATCTATTCATGTCCTGAGTGCCTTCTTGTCCACAGAGATCTA 315
DB 43 PheSerLeuMetProCysAlaLeuValGlnPheThrLeuLeuPheValHisArgAspLeu 62
QY 316 GCCAAGATAAACCGCTATCATTTATTTATGATCTTAATCTCTTGGGACCTGTTATCAGA 375
DB 63 SerArgAspArgProLeuAlaLeuLeuMetHisLeuLeuGlnLeuGlyProLeuTyrArg 82
QY 376 TGTGTTGAGCGCATGATTAAGTACTCTACATGCTGGGAAGAAGAGACGAGAGAGCC 435
DB 83 CysCysGlnValPheCysIleTyrCys-----GlnSerAspGlnAsnGluGluPro 99
QY 436 TATGTCAGCTCACCCGNAAGAAG---ATGCTAATAGTGGGAGGAGGCTGCTGATAGAA 492
DB 100 TyrValSerIleThrLysLysArgGlnMetProLysAspGlyLeuSerGluGluValGlu 119
QY 493 TGGGAGGTGGGCCACTCCATCCGACCTGCTATGCTGACCGCAATGCTCAAAAGATG 552
DB 120 LysGluValGlyGlnAlaGluGlyLysLeuIleThrHisArgSerAlaPheSerArgAla 139
QY 553 TCACAGATCCAAAGCCTTCTCTGGGCTCAGTGCCCGACGCTGACCTATCAGCTTATGTAGC 612
DB 140 SerValIleGlnAlaPheLeuGlySerAlaProGlnLeuThrLeuGlnLeuTyrIleThr 159
QY 613 CTGATCTCTGCAGAGGTTCCCTCGGTAGAGTTGTGCTAATGGTATTTTCCCTGCTATCT 672

```

```

DB 160 ValLeuGluGlnAsnIleThrThrGlyArgCysPheIleMetThrLeuSerLeuSer 179
QY 673 GTCACCTATGTTGGGCGCCACCTTTTGGCAATATGTTGGTATCCAGATCAAGTACGATCTAC 732
DB 180 IleValTyrGlyAlaLeuArgCysAsnIleLeuAlaIleLysIleLysTyrAspGluTyr 199
QY 733 AAGATTGCGCTTGGCGCACCTAGAGTCTCTGTCATCACCATCTGGCGACATTTGGAGATC 792
DB 200 GluValLysValLysProLeuAlaTyrValCysIlePheLeuTrpArgSerPheGluIle 219
QY 793 ACTTCCCGCTCTCTGATTCTGCTGCTCTTCTCAGCCACTTTTGAATTTGAAGCTCTGCCC 852
DB 220 AlaThrArgValIleValLeuValLeuPheThrSerValLeuLysIleTrpValAla 239
QY 853 TTCCTAGTGTCAACTCTCTCTGATCATCTCTTTCAGCCCTGATTAAGTTTGGAGAGT 912
DB 240 ValIleLeuValAsnPhePheSerPhePheLeuTyrProTrpIleValPheTrpCysSer 259
QY 913 GGTGCCCGACATGCCCAATTAACATTGAGAAAACCTTCAGCCGCTCGGCACTCTGGTGGTC 972
DB 260 GlySerProPheProGluAsnIleGluLysAlaLeuSerArgValGlyThrThrIleVal 279
QY 973 CTGATTTTCAGTCACCATCTCTATGCTGGCATCAACTTCTCTTGTGCTGCTGCTTTCAG 1032
DB 280 LeuCysPheLeuThrLeuLeuTyrAlaGlyIleAsnMetPheCysTrpSerAlaValGln 299
QY 1033 TTGAGGTTGGCAGACAGAGATCTCTGTCACAAAGGCGAGAACTGGGACATATGGCCCTG 1092
DB 300 LeuLysIleAspAsnProGluLeuIleSerLysSerGlnAsnTrpTyrArgLeuLeuIle 319
QY 1093 CACTATAGTGTGAGGTTGGTAGAAGATGTGATCATGCTTGTGCTTGTGCTTCTTGA 1152
DB 320 TyrTyrMetThrArgPheIleGluAsnSerValLeuLeuLeuLeuTyrTyrPhePheLys 339
QY 1153 GTGAAAGTGTACTGAATCTCTCTGATTCCTTGTGCTGCTGCTGCTGCTTATTCCTTAT 1212
DB 340 ThrAspIleTyrMetTyrValCysAlaProLeuLeuIleLeuGlnLeuLeuIleGlyTyr 359
QY 1213 CTGATTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
DB 360 CysThrGlyIleLeuPheMetLeuValPheTyrGlnPhePheHisProCysLysLysLeu 379
QY 1273 TTCACCATATAGTAGAC-----TACCTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 380 PheSerSerValSerGluSerPheArgAlaLeuLeuA-gCysAlaCysTrpSerSer 399
QY 1321 CACCTCGGACGAGGTTGAGAACTCAGAGCCA 1353
DB 400 -----LeuArgArgLysSerSerGluPro 407

```

RESULT 2

```

Q9H6D3 ID Q9H6D3 PRELIMINARY; PRT; 395 AA.
AC Q9H6D3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CDNA: FLJ22371 fis, clone HRC06680 (Hypothetical 44.7 kDa protein)
DS (Similar to hypothetical protein FLJ10307).
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK026024; BAB15326.1; -;
DR EMBL; BC013379; AAH13379.1; -;
DR EMBL; BC028564; AAH28564.1; -;

KW Hypothetical protein.

SQ SEQUENCE 395 AA; 44654 MW; 23199BAEEA6964C6 CRC64;

Alignment Scores:

Pred. No.:	1.87e-06	Length:	395
Score:	158.50	Matches:	92
Percent Similarity:	38.33%	Conservative:	82
Best Local Similarity:	20.26%	Mismatches:	161
Query Match:	6.23%	Indels:	119
DB:	4	Gaps:	21

US-09-768-781-2 (1-1389) x Q9H6D3 (1-395)

```

QY 91 TCTCTGAGAGATGTCATCCGTGGAGCAACCCCGATTACTTTTCCATTAGCATC 150
Db 8 AlaLeuLeuArgAspLeuValLeuGly----- 16
QY 151 CTTTCTCCACCTTTTGTACTGTGGGAGGCTGCATCTCTGTGTACGTGGTAGAATC 210
Db 17 ValLeuGlyThrAlaAlaPheLeuLeuAspLeuGlyThrAspLeuTrpAlaAlaValGln 36
QY 211 TATCGAAAGATAGTGAACCTTACTGTGATGACATACACCTTTTCTTTCTTATGTTTCA 270
Db 37 TyrAlaLeuGlyArgTyrLeuTrpAlaAlaLeuValLeuAlaLeuLeuGlyLeuAla 56
QY 271 TCCATTATGTCAGTTG---ACCCTCATTTTGTCCACAGAGAT-----CTAGCC 318
Db 57 SerValAlaLeuGlnLeuPheSerTrpLeuTrpLeuArgAlaAspProAlaGlyLeuHis 76
QY 319 AAAGATAAACCG-----CTATCATTTATTTATGATCTAATCTCTCTGGACCT 366
Db 77 GlySerGlnProProArgArgCysLeuAlaLeu---LeuHisLeuLeuGlnLeuGlyTyr 95
QY 367 GTTATCAGATGTTTGGAGGCCATGATTAAGTACTCACACTGTGGAGAAAGAGGAGCAG 426
Db 96 LeuTyrArgCysValGlnLeuLeuArgGlnGlyLeuLeuValTrp-----GlnGln 112
QY 427 GAGGAGCCCTATGTCAGCCCTCACCCGAAAGAGATGCTAATAGATGGCGAGGAGTGTG 486
Db 113 GluGluPro----- 115
QY 487 ATAGATGGGAGTGGGCGCACTCCATCCGACCCCTGGCTATGACCCCAATGCTACAA 546
Db 116 SerGluPheAspLeuAlaTyrAla---AspPheLeuAlaLeuAsp-----IleSer 131
QY 547 CGTATGTCACAGATCCAAAGCCTCTCTGGGCTCAGTCCCGCCAGCTACCTATCAGCTCTAT 606
Db 132 MetLeuArgLeuPheGluThrPheLeuGluThrAlaProGlnLeuThrLeuValLeuAla 151
QY 607 GTGAGCCTGATCTCTGACAGAGGTTCCCTGGGTAGAGTTGTGTAATGGTATTTTCCCTG 666
Db 152 IleMetLeuGlnSerGlyArgAlaGluTyrTyrGlnTrpValGlyIleCysThrSerPhe 171
QY 667 GTATCTGTCACCTATGGGCGCCACCTTTTGCATATATGTTGGCTATCCAGATCAAGTACGAT 726
Db 172 LeuGlyIleSerTrpAlaLeuLeu----- 179
QY 727 GACTAC-----AAGATTCGCTTGGGCCACTAGAA 756
Db 180 AspTyrHisArgAlaLeuArgThrCysLeuProSerLysProLeuLeuGlyLeuGlySer 199
QY 757 GTCCTCTGCATCACCATCTGGCGGAGATTGGAGATCACTTCCCGCCTCTGATTCTGGTG 816

```

```

Db 200 SerValIleTyrPheLeuTrpAsnLeuLeuLeuLeuTrpProArgValLeuAlaValAla 219
QY 817 CTCTTCTCAGCCACTTTG---AAATTGAAGCTGTGCCCTTCCTAGTGTCAACTTCTCTG 873
Db 220 LeuPheSerAlaLeuPheProSerTyrValAlaLeuHisPheLeuGlyLeuTrpLeuVal 239
QY 874 ATCATCCTCTTTAGCCCTGGATTAAAGTTCTGGAGAAAGTGTGTGCCAG---ATGCCCAAT 930
Db 240 LeuLeuLeu-----TrpVal-----TrpLeuGlnGlyThrAspPheMetProasp 254
QY 931 AACATTGAGAAAACTTCAGCCGGTGGCAGCTCTGTGTGGTCTGATTTCAGTCCACCATC 990
Db 255 ProSerSerGluTrpLeuTyrArgValThrVal-----AlaThrIle 268
QY 991 CTCTATGCTGGCATCAACTTCTCTGTGTGTGTCAGCTTTGCGAGTTGCGCAGACAGA 1050
Db 269 LeuTyr-----PheSerTrpPheAsn----- 275
QY 1051 GATCTCGTCGACAAAGGGCAGACATATGGCCCTGCACATATAGTGTGAGGTTG 1110
Db 276 -----ValAlaGluGlyArgThrArgGlyArgAlaIleIleHisPheAlaPheLeuLeu 293
QY 1111 GTAGAGAAATGTATCATGTCTTGTGTTTAAAGTTCTTTTGGAGTGAAGTGTACTGAAT 1170
Db 294 SerAspSerIleLeuLeuValAlaThrTrpValThrHisSer-----Ser 308
QY 1171 TACTGTATCTCTTGTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1215
Db 309 TrpLeuProSerGlyIleProLeuGlnLeuTrpLeuProValGlyCysGlyCysPhePhe 328
QY 1216 ATTTCCATGTCCTTATGTCCTCTCTTCTTCCAGTACTTGTGCTGCTGCTGCTGCTGCTGCT 1275
Db 329 LeuGlyLeuAlaLeuArgLeuValTyrTyrHisTrpLeuHisPro----- 343
QY 1276 ACCATAATGTAGTACTACCTCCATTGTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1335
Db 344 -----SerCysCysTrpLysProAspProAspGln 353
QY 1336 GTTGAGAACTCAGAGCCACCTTTGTAGACTGAAGCAAGGCAA 1377
Db 354 ValAspGlyAlaArgSerLeuLeuSerProGluGlyTyrGln 367

```

RESULT 3

```

Q96P28 ID Q96P28 PRELIMINARY; PRT; 505 AA.
AC Q96P28;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE KIAA1889 protein (Fragment).
GN KIAA1889.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21456161; PubMed=11572484;
RA Nagase T., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XXI.
RT The complete sequences of 60 new cDNA clones from brain which code for
RT large proteins."
RL DNA Res. 8:179-187(2001).
DR EMBL; AB067476; BAB67782.1; -.
FT NON_TER
SQ SEQUENCE 505 AA; 56911 MW; D764B96BC29B3623 CRC64;

```

Alignment Scores:

Pred. No.:	8.48e-05	Length:	505
Score:	141.50	Matches:	65
Percent Similarity:	42.41%	Conservative:	69

Pred. No.: 0.00128 Length: 362
 Score: 129.00 Matches: 58
 Percent Similarity: 40.98% Conservative: 42
 Best Local Similarity: 23.77% Mismatches: 100
 Query Match: 5.07% Indels: 44
 DB: 11 Gaps: 9

US-09-768-781-2 (1-1389) x Q8RI18 (1-362)

QY 559 ATCAAGCTTCCTGGGCTCAGTCCAGCTGACCTATCAGCTCTATGTGAGCTGATC 618
 DB 14 LeuGluThrPheLeuGluSerAlaProGlnLeuValLeuGlnLeuCysIleMetIleGln 33
 QY 619 TCTCAGAGGTTCCCTCGGTGAGATTGTGTAATGGTATTTTCCCTGGTATCTCTGACC 678
 DB 34 LysAsnSerAlaGluThrLeuProCysValSerValThrSerLeuMetSerLeuAla 53
 QY 679 TATGGGGCCACCTTTGCAATATGTTGCTATCCAGATCAAGTACAGATGATC 738
 DB 54 TrpValLeuAlaSerTyrHisLeuLeuLeuArgAspSerArgAspAspLysLys 71
 QY 739 CGCTTGGGGCCTAGAGATCTCTGTCATCACCATCTCGGGCAGATTGGAGATCACTCC 798
 DB 72 SerMetSerTyrArgGlyAlaLeuIleHisLeuPheTrpArgLeuPheThrIleSerSer 91
 QY 799 CGCTCTCTGATCTGCTCTCTCAGCCACTTTGAAATGAGGCTGTGCCCTTCCTA 858
 DB 92 ArgValIleSerPheAlaLeuPheAlaSerIlePheGlnLeuTyrPheGlyIlePheVal 111
 QY 859 GTGCTCAACTCTCGATCATCTCTTTGAGCCCTGGATTAAGTTCTGGAGAGTGGTCC 918
 DB 112 ValValHisTrpCysAlaMetAlaPhe-----TrpIleIleHisGlyThrAspPhe 129
 QY 919 CAGATGCCCAATACATTGAGAAAACCTTCAGCCGGTCCGCACTCTGCTGCTGATT 978
 DB 130 CysMetSerTyrTrpGluGluIleLeuPheAsnMetVal----- 142
 QY 979 TCAGTCACTCTCTATGCTGGCATCAACTTCTCTGCTGTCAGCTTTGCGAGTTGAGG 1038
 DB 143 ---ValGlyIleValTyr-----IlePheCysTrpPheAsn----- 153
 QY 1039 TTGGCAGACAGATCTGTCACAAAGGGCAGAACTGGGACATATATGGCTGCACTAT 1098
 DB 154 -----ValLysGluGlyArgThrArgTyrArgMetPheAlaTyrTyr 167
 QY 1099 AGTGTAGTGTGTAGAAATGTGATCATGTCTTGGTTTAAAGTTC----- 1146
 DB 168 ThrIleValLeuThrGluAsnAlaAlaLeuThrPheLeuTyrPheTyrArgAsnPro 187
 QY 1147 -----TTGGAGTGAAGTGTACTGAATTACTGCTCATCTCTGATTGCC 1191
 DB 188 GluSerThrAspSerTyrAlaValProAlaLeuCys-----Cys----- 200
 QY 1192 TTGCAGCTCAATTATGCTATCTGATTTCCATTTGCTTCATGCTCTTCTTCCAGTAC 1251
 DB 201 -----ValPheValSerPheValAlaGlyIleThrLeuMetLeuTyrTyrGlyVal 218
 QY 1252 TTGCATCATTTG 1263
 DB 219 LeuHisProMet 222

RESULT 7

Q9SAK8 PRELIMINARY; PRT; 783 AA.
 AC Q9SAK8
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE T8K14.18 protein.
 GN T8K14.18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
 Li J., Kremenetskaia I., Luros J., Lee J.M., Gonzalez A., Altafi H.,
 Araujo R., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,
 Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,
 Davis R.W., Ecker J.R., Federpiehl N.A., Theologis A.,
 RT "Arabidopsis thaliana chromosome 1 BAC T8K14 sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007202; AAD30236.1; -
 DR InterPro; IPR000676; NaH Exchng.
 DR Pfam; PF00999; Na.H Exchanger.1.
 SQ SEQUENCE 783 AA; 88203 MW; 0A6846910C381648 CRC64;

Alignment Scores:

Pred. No.: 0.00771 Length: 783
 Score: 121.50 Matches: 89
 Percent Similarity: 34.58% Conservative: 68
 Best Local Similarity: 19.60% Mismatches: 166
 Query Match: 4.78% Indels: 131
 DB: 10 Gaps: 19

US-09-768-781-2 (1-1389) x Q9SAK8 (1-783)

QY 256 TTCCTTATGTTTTCATCATTCATGTCAGTTGACCTCATTTTGTTCACAGATCTA 315
 DB 16 PheAsnProLeuAsnThrMetPheIleGlnMetAlaCysIleLeuValPheSerGlnLeu 35
 QY 316 -----GCCAAAGATAAACCGCTATCATTTATGATCTA 351
 DB 36 PheTyrLeuLeuLeuLysProCysGlyGlnAlaGlyProValAlaGlnIleLeuAlaGly 55
 QY 352 ATCTCTTGGGACCTGTTATCAGATGTTTGGAGGCCATGATTAACTCCTCACACTGTGG 411
 DB 56 IleValLeuSerProValLeuLeuSerArgIleProLysValLysGluPheLeuGln 75
 QY 412 AAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 465
 DB 76 LysAsnAlaAlaAspTyrTyrSerPhePheSerPheAlaLeuArgThrSerPheMetPhe 95
 QY 466 ATAGATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 525
 DB 96 LeuIleGlyLeuGluValAspLeuHisPhe----- 105
 QY 526 ATGCACCGCAATGCTTACAAACGATGTCACAGATC-----CAAGCCTTCTCTG----- 573
 DB 106 MetArgArgAsn---PheLysLysAlaAlaValIleThrLeuSerSerPheValValSer 124
 QY 574 GGCTCAGTGGCCCCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCTGCAGAGGTTCC 633
 DB 125 GlyLeuLeuSerPheAlaSerLeuMetLeuPheIleProLeuPheGlyIleLysGluAsp 144
 QY 634 CTGGGTAGAGTTGCTTAATGGTATTTTCCCTGATCTGTACCTATGGGGCCACCTT 693
 DB 145 TyrPheThrPhePheLeuValLeuValThrLeuSerAsnThrAlaSerProValVal 164
 QY 694 TGCAATATGTTGGCTATCCAGATCAAGTACGATCACTACAAAGATT-----CGCCTT 744
 DB 165 ValArgSerIleAla-----AspTrpLysLeuAsnThrCysGluIle 178
 QY 745 GGGCCACTAGAGTCTCTTCATCACCATCTGGGGGACATTTGGAGATCATCTTCCCGCCTC 804
 DB 179 GlyArgLeuThrIleSerCysAlaLeuPheIleGluLeuThrAsnValValLeuTyrThr 198
 QY 805 CTGATTTCTGGTCTCTCTCAGCCACTTTGAAATGAGGCTGTGCCCTTCTCTAGTGTCTC 864
 DB 199 IleIleMetAlaPheIleSerGlyThrIleIleLeuGluLeuPheLeuPheLeuAla 218
 QY 865 AACTTCTCTGATCATCTCTTTGAG-----CCCTGGATTAACTTCTGGAGAGT 912

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
RT clone:B1060H01";
RL
RN Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[2]

```

RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT clone:OSNBNB0036G09.";
RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RL
DR EMBL; AP003560; BAB68079.1; -.
DR ENBL; AP003309; BAB89942.1; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR002916; Ferric_reduct.
DR Pfam; PF00036; ehfand; 1.
DR Pfam; PF01794; Ferric_reduct; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
SQ SEQUENCE 943 AA; 106257 MW; 6B5EDE1339D28657 CRC64;

```

Alignment Scores:		
Pred. No.:	0.0094	943
Score:	120.50	Matches: 78
Percent Similarity:	36.44%	Conservative: 59
Best Local Similarity:	20.74%	Mismatches: 144
Query Match:	4.74%	Indels: 95
DB:	10	Gaps: 15

US-09-768-781-2 (1-1389) x Q942C7 (1-943)

310	QY	GATCTAGCCAAAGATAAACCGCTTATCATTTATTCATCTAATCTCTTTGGGACCTGTT	369
256	Db	AspAenSerPheAepSerArgLeuGlnIlePhePheGluMetValAspLysAsnAlaAep	275
370	QY	ATCAGATGTTTGGAGGCCATGATTAAAGTACTCTC-----ACA	405
276	Db	GlyArgIleThrGluAlaGluValLysGluIleLeuMetLeuSerAlaSerAlaAsnLys	295
406	QY	CTGTGGAAAGAGGAGGAGCGCCCTATGTTCAGCGCTCACCCGAAGAAGATGCTTA	465
296	Db	LeuSerArgLeuLysGluGlnAlaGluGluTyrAlaAlaLeuIleMetGluGluLeuAep	315
466	QY	ATAGATGGCGAGGAGGTGCTGATAGATAGGAGGTGGCCACTCCATCGGACCCCTGGCT	525
316	Db	ProGluGlyLeuGlyTyrIleGluLeuTrpGln-----LeuGluThrLeuLeu	331
526	QY	ATGCACCGCAATGCCCTACAAACGTATGTTCACAGATC-----CAAGCCTTC	570
332	Db	LeuGlnLysAepThrTyrMetAsnTyrSerGlnAlaLeuSerTyrThrSerGlnAlaLeu	351
571	QY	CTGGGCTCAGTGGCCCGAGCTGACCTATCATCAGCTCTATGTGAGCGCTGATCTCTGCAGAGGTT	630
352	Db	SerGlnAsnLeuAlaGlyLeuArgLysLysSerSerIleArgLysIleSer-----	368
631	QY	CCCTCGGGTAGAGTGTGCTAAATGGTATTTCCTCGTATCTCTCACTATGGGGCCACC	690
369	Db	-----ThrSerLeuSerTyr-----	373
691	QY	CTTTGCAATATGTTGGCTATCCAGATCAAGTAGCAGTACAGATTCACCAAGATTCCCTTTGGGCA	750
374	Db	-----TyrPheGluAspAsnTrpLysArgLeuTrpVal	384
751	QY	CTAGNAGTCCTCTGCATCACCATC-----TGGCGGACATTTGGAGTC	792
385	Db	Leu-----AlaLeuTrpIleGlyIleMetAlaGlyLeuPheThrTrpLysPheMetGlnTyr	403
793	QY	ACTTCCCGCTCTCTGATTCTGGTGCTC---TTCTCAGCCCATCTTGAATTGAAGCGTGTG	849
404	Db	ArgAsnArgTyrValPheAspValMetGlyTyrCysValThrThrAlaLysGlyAlaAla	423
850	QY	CCCTTCCTAGTGTCAACTCTCTGATCATCTCTTTAG-----CCC	891
424	Db	GluThrLeuLysLeuAsnMetAlaIleLeuLeuProValCysArgAsnThrIleThr	443

Qy	892	TGGATTAAAGTTCTGGAGAAGTGGTGCCAGATAGCCCAATAACATTGAGAAAAACTTCAGC	951
Db	444	TrpLeuArgSerThrArgAlaAlaAlaArgAlaLeuProPheAspAsnIleAsnPheHis	463
Qy	952	CGGGTCGCACACTCTGGTGGTCTGATTTCAGTCACCATCTCTATGCTGGCATCAACTTC	1011
Db	464	LysThrIleAlaAlaIleValValGlyIle---IleLeuHisAlaGlyAsnHisLeu	482
Qy	1012	TCTTGGTGGTCAGCTTTGCAGTTGAGGTGGCAGACAGAGACTCTCGTCGACAAAGGCGAG	1071
Db	483	ValCysAspPheProArgLeuIleLysSerSerAspGluLysTyrAlaProLeuGlyGln	502
Qy	1072	AACTGGGACATATGGCGCTGCACATAT-----AGTGTGAGGTGGTAGAG	1116
Db	503	TyrPheGlyGluIleLysProThrTyrPheThrLeuValLysGlyValGluGlyIleThr	522
Qy	1117	AATGTGATCGTCTGCTTCGTTT-----1140	
Db	523	GlyValIleMetValValCysMetIleIleAlaPheThrLeuAlaThrArgTrpPheArg	542
Qy	1141	-----AAGTTCTTTGGAGTGAAAGTGTTACTG	1167
Db	543	ArgSerLeuValLysLeuProArgProPheAspLysLeuThrGlyPheAsnAlaPheTrp	562
Qy	1168	AATTACTGTCATCTCTTG-----ATTGCCTTGCAGCTCATATT-----	1206
Db	563	---TyrSerHisHisLeuPheIleIleValTyrIleAlaLeuIleValHisGlyGluCys	581
Qy	1207	GCTTATCTGATTTCATCGGCTCATGCTCCCTTTTCTTCCACTACTTG	1254
Db	582	LeuTyrIleuIleHisValTrpTyrArgArgThrThrTrpMetTyrLeu	597
RESULT 11			
ID	Q9B8Z8	PRELIMINARY; PRT; 424 AA.	
AC	Q9B8Z8;		
DT	01-JUN-2001	(TrEMBLrel. 17, Created)	
DT	01-JUN-2002	(TrEMBLrel. 17, Last sequence update)	
DT	01-MAR-2002	(TrEMBLrel. 20, Last annotation update)	
DE	NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).		
GN	NAD4.		
OS	Schistosoma japonicum (Blood fluke).		
OC	Mitochondrion.		
OC	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;		
OC	Schistosomatidae; Schistosomatidae; Schistosoma.		
NCBI_TaxID	6182;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=ANHUI.		
RX	MEDLINE=20349913; PubMed=1089225;		
RA	Le T.H., Blair D., Agatsuna T., Humair P.F., Campbell N.J.,		
RA	Iwagami M., Littlewood D.T., Peacock B., Johnson D.A., Bartley J.,		
RA	Rollinson D., Herniou E.A., Marchal D.S., McManus D.P.;		
RT	"Phylogenies inferred from mitochondrial gene orders-a cautionary tale		
RT	from the parasitic flatworms";		
CC	Mol. Biol. Evol. 17:1123-1125(2000).		
CC	-I- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.		
DR	EMBL; AF215860; AAG13138.2; .		
DR	InterPro; IPR003918; NADHub OxRed4.		
DR	InterPro; IPR001750; Oxidored_q1.		
DR	Pfam; PF00361; oxidored_q1; 1.		
DR	PRINTS; PR01437; NUOXORDTASE4.		
SK	Mitochondrion; NAD; Oxidoreductase; Ubiquinone		
SEQUENCE	424 AA; 47741 MW; 377C53E611F57D72 CRC64;		

Alignment Scores:		
Pred. No.:	0.0108	Length:
Score:	119.50	Matches:
Percent Similarity:	36.3%	Conservative:
Best Local Similarity:	19.70%	Mismatches:
Query Match:	4.70%	Indels:
DB:	8	Gaps:
		424
		51
		77
		155
		139
		20


```
Db 72 ArgProLeuSerTyrLysGlyAlaValAlaGlnValLeuTrpHisLeuPheSerIleAla 91
Qy 796 TCCGCGCTCCTGATCTGGTCTCTTCAGCGCATTTGAAATTAAGAGCTGTGCGCTTC 855
Db 92 AlaArgGlyLeuAlaPheAlaLeuPheAlaSerValTyrLysLeuTyrPheGlyIlePhe 111
Qy 856 CTAGTGTCTCAACTCTCTGATCATCTCTTTGAGCCCTGGATTAAGTTCTGGAGAGTGTG 915
Db 112 IleValAlaHisTrpCysValMetThrPhe-----TrpVal----- 123
Qy 916 GCCAGATGCCCAATACATTGAGAAAACCTTCACCGGGTCGGCACTCTGGTGTCTCTG 975
Db 124 -----IleGlnGlyGluThrAspPheCysMetSerLysTrpGluGluIleIleTyrAsn 141
Qy 976 ATTTCACTCACCATCTCTATGCTGGCATCACTCTCTGCTGTGTGCTGAGCTTTCAGTTG 1035
Db 142 MetValValGlyIleIleTyr-----IlePheCysTrpPheAsn----- 154
Qy 1036 AGTTGGCAGACAGATCTCTGTCACAAAGGGCAGAACTGGGGACATATGGGCTGCAC 1095
Db 155 -----ValysGluGlyArgSerArgArgMetThrLeuTyr 167
Qy 1096 TATAGTGTGAGTGTGTAGAGAATGTGATGCTGTGGTCTTTTAAAGTTCTTTGGAGTG 1155
Db 168 HisCysIleValLeuLeuGluAla-----AlaLeuThrGlyPheTrpTyrSerSer 185
Qy 1156 AAAGTGTACTGAATTAATCTGATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1212
Db 186 ArgAsnPheSerThrAspPheTyrSerLeuIleMetValCysValValAlaSerSerPhe 205
Qy 1213 CTGATTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 206 AlaLeuGlyIlePhePheMetCysValTyrTyrCysLeuLeuHisPro 221

RESULT 13
Q94UX1 PRELIMINARY; PRT; 424 AA.
AC Q94UX1
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE NAD4 dehydrogenase subunit 4.
GN NAD4
OS Schistosoma japonicum (Blood fluke).
OG Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20349913; PubMed=10899225;
RA Le T.H., Blair D., Agateuma T., Humair P.F., Campbell N.J.,
RA Iwagami M., Littlewood D.T., Peacock B., Johnston D.A., Bartley J.,
RA Rollinson D., Harniou E.A., Zaranga D.S., McManus D.P.;
RT "Phylogenies inferred from mitochondrial gene orders-a cautionary tale
RT from the parasitic flatworms."
RL Mol. Biol. Evol. 17:1123-1125(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Le T.H., Blair D., McManus D.P.;
RT "Revisiting limited genetic variation within Schistosoma japonicum
RT populations."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL: AF412213; AAL12155.1;
DR InterPro: IPR001750; Oxidored_q1.
DR Pfam: PF00361; Oxidored_q1; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 424 AA; 47725 MW; 1BC880B001235E31 CRC64;

Alignment Scores: 0.0135 Length: 424
Pred. No.:
```

```
Score: 118.50 Matches: 91
Percent Similarity: 36.36% Conservative: 77
Best Local Similarity: 19.70% Mismatches: 155
Query Match: 4.68% Indels: 139
DB: 8 Gaps: 20
US-09-768-781-2 (1-1389) x Q94UX1 (1-424)
Qy 133 ACTTTTCCATTAGCATCTTTTCTCCACCTTTTGTACTGTGGGAGGCTGCATCTGCT 192
Db 18 ThrPheIlePheLeuLeuIlePheTyrSerIleIleTrpValSerAspSerMetVal 37
Qy 193 TTGTACATGGTTAGTAATCTATCGAAAGAAATAGTGAACCTTACTGGATGACATACACCTTT 252
Db 38 MetValGlyValLysTyrTyrLeuCysAspGlyLeuValIleAspThr----- 54
Qy 253 TCTTCTTTTATGTTTTCATTCATATGTCAGTTGAC-----CTCATTTTGTCCACAGA 309
Db 55 -----LeuSerCysLeuMetIlePheLeuThrSerIleIleTrpLeuValLeu 70
Qy 310 GATCTAGCCAAAGATAAACCCGCTATCATTTATTCATCTAATCTCTTGGACCTGTT 369
Db 71 TrpLeuValGlySerLysAspIleValLeuPhe----- 81
Qy 370 ATCAGATGTTGGAGGCCATGATTAAGTAC-----CTCACACTGTGGAAG 414
Db 82 IleSerValPheSerAlaMetLeuThrTyrValValSerAsnSerLeuValPheTrpPhe 101
Qy 415 AAAGAGGAGCAGGAGGCCCTATGTCAGCTTACCCGAAAGAGATGCTAATAGATGC 474
Db 102 PheTyrGluLeuSer-----IleIleSerAlaLeuTyrMetLeuIleValGly 117
Qy 475 -----GAGGAGTGTGATAGATGGAGGTGGCCACTCCATCCGACCCCTG 522
Db 118 SerProTyrProGluArgTyrIleSerSerTrpTyrPheGlyTyrIle----- 134
Qy 523 GCTATGCACCGCAATGCCTACAAAGTATGTACAGATCCAAAGCCTTCTCGGCTCAGTG 582
Db 135 -----LeuLeuSerSerVal 139
Qy 583 CCCAGCTGACCTATCAGCTCTATGTAGCCTGATCTCTGCAGAGGTTCCTCGGTGAG 642
Db 140 ProLeuLeuLeuGlyIleCysPheIleGlyLeuAsnSerGlySerPhe-----Asn 156
Qy 643 GTTGTGCTAATGTTATTTTCCCTGCTATCTGTACCTATGTCAGTGGGCGCCATGTCATATG 702
Db 157 ValIleLeuTyrAspLysGlyAspMetCysAspSerTyrGlyAlaPheLeuLeuIle 176
Qy 703 TTG----- 705
Db 177 ValMetPheLeuThrLysIleProValPheProPheHisGlyTrpLeuProLeuValHis 196
Qy 706 -----GCTATCCAGATCAAGTACGATGACTACAAGATTCGCTTGGGCCA 750
Db 197 AlaGluAlaSerSerProValSerIleIleLeuSerGlyTyrIleMetLysLeuGlyLeu 216
Qy 751 CTAGAAGTCTCTGTCATCACCATCTGCGGACATTTGGAGATCATTCCCGCTCTCGATT 810
Db 217 ValGlyLeuValArgLeuCysGlyTrpLeuLeuIleAsp-----Tyr 230
Qy 811 CTGTGTCTTCTCAGCCACTTTGAAATTAAGGTGAGGCTGTGCTCTCTAGTCTCACTTC 870
Db 231 IleTyrTyrPheSerThrPheLeuLeuCysTyrSerValValPheLeuVal----- 247
Qy 871 CTGATCATCTCTTTGAGCCCTGGATTAAGTTCTGGAAGAGGTGGTGGCCAGATGCCCAAT 930
Db 248 ---AlaAlaValPheGluCysAspSerLysArgTrpLeuAlaTyrLeuSerLeuSerHis 266
Qy 931 AACATTGAGAAAAAATTTCAGCGGCTCGGCACCTCTGGTGGTCTGCTG----- 975
Db 267 IleLeu-----IleGlyValCysIleLeuLeuThrSerThrTyrCys 280
Qy 976 -----ATTTCAGTCACCATCTCTAT 996
```

```

Db 281 GlyAspTyrLeuAlaPheIleTyrCysLeuGlyHisGlyLeuSerValAlaLeuPhe 300
Qy 997 GCTGGCATCAACTCTCTGCTGGTGCAGCTTGGTGGTGGGAGGCTGCATCTGCT 1056
Db 301 MetIleIle-----TrpPheGlyTyrGlu----- 308
Qy 1057 GTCGACAAAGGCGAGAACTGGGACATATGGGCGCTGCACATAGTGTGAGGTTGGTAGAG 1116
Db 309 IleserGlySerArgAsnTrpGlyIleLeuValIlePheGlyGlyGlyLeuMet 328
Qy 1117 AATGTGATCATGCTGCTGGTGT-----AAGTTC 1146
Db 329 HisPheIleMetGlyPheValPheLeuAsnValCysGlyPheProAlaLeuGlnPhe 348
Qy 1147 TTGGGA---GTGAAAGTGTACTGAATTAC-----TGTCACTTCCTTGCATTCGCTGCAG 1197
Db 349 PheGlyLeuLeuTrpLeuValIleAsnTyrIleThrLeuGlyAspIleIleSerLeu 368
Qy 1198 CTCATTATTCCTATCTGATTTC-----ATTGGCTTCATGCTCTTCTTCCAG 1248
Db 369 LeuValSerIleTyrIlePheSerGlySerIleIleGlyPheIleIleTyrGlyLeuVal 388
Qy 1249 TACTGTCATTCGCTCCTCTTCCACCAAT-----GTAGTAGACTACTCCAT 1302
Db 389 IleCysSerProIleAsnThrSerTyrGlyTyrSerGlyGlyLeuAspAsnPheLeuPhe 408
Qy 1303 TGTGTC 1308
Db 409 CysIle 410

RESULT 14
O94UX2 PRELIMINARY; PRT; 424 AA.
AC Q94UX2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN NAD4.
OS Schistosoma japonicum (Blood fluke).
OC Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2034913; PubMed=10889225;
RA Le T.H., Blair D., Agatsuma T., Humair P.F., Campbell N.J.,
RA Iwagami M., Littlewood D.T., Peacock B., Johnston D.A., Bartley J.,
RA Rollinson D., Herniou E.A., Zariwaga D.S., McManus D.P.;
RT "Phylogenies inferred from mitochondrial gene orders-a cautionary tale
RT from the parasitic flatworms."
RL Mol. Biol. Evol. 17:1123-1125(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Le T.H., Blair D., McManus D.P.;
RT "Revisiting limited genetic variation within Schistosoma japonicum
RT populations."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL; AF121211; AAL12147.1;
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 424 AA; 47669 MW; D197589A496177B8 CRC64;

```

```

Alignment Scores:
Pred. No.: 0.021 Length: 424
Score: 116.50 Matches: 90
Percent Similarity: 36.36% Conservative: 78
Best Local Similarity: 19.48% Mismatches: 155
Query Match: 139 Indels: 139

```

```

DB: 8 Gaps: 20
US-09-768-781-2 (1-1389) x Q94UX2 (1-424)
Qy 133 ACITTTCCATTAGCATCTTTTCTCCACCTTTTGTACTGTGGGAGGCTGCATCTGCT 192
Db 18 ThrPheIlePheLeuIlePheTyrSerIleIleTrpValSerAspSerMetVal 37
Qy 193 TTGTACATGTTAAATCTATCGAAAGAAATAGTGAATCTACTGGATGACATACACCTTT 252
Db 38 MetValGlyValIlyTyrTyrLeuCysAspGlyLeuValIleIleAspThr----- 54
Qy 253 TCTTTCTTTATGTTTTCATCATTCATGTCAGTGTGACC-----CTCATTTTGTCCACAGA 309
Db 55 -----LeuSerCysLeuMetIlePheLeuThrSerIleIleTrpLeuValLeu 70
Qy 310 GATCTAGCCAAAGATAAACCGCTATCATTTATGATCTAATCTCTTGGGACCTGTT 369
Db 71 TrpLeuValGlySerIlyAspIleValLeuPhe----- 81
Qy 370 ATCAGATGTTTGGAGGCGCATGATTAAAGTAC-----CTCACACTGTGGAAG 414
Db 82 IleserValPheSerAlaMetLeuThrTyrValValSerAsnSerLeuValPheTrpPhe 101
Qy 415 AAAGAGGAGCAGGAGGAGCCCTATGTCAGCCTCACCCGAAAGAGATGCTAATAGATGCG 474
Db 102 PheTyrGluLeuSer-----IleIleSerAlaLeuTyrMetLeuIleValGly 117
Qy 475 -----GAGAGGTGCTGATAGATGAGAGGTGGGCGCCTCCATCGGACCCCTG 522
Db 118 SerProTyrProGluArgTyrIleSerSerTyrTyrPheGlyTyrIle----- 134
Qy 523 GCTATGCACCCCAATGCTACAAACGTATGTCACAGATCCAAAGCCTTCTCGGCTCAGTG 582
Db 135 -----LeuLeuSerSerVal 139
Qy 583 CCCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCTGCAGAGGTTCCCTCGGTAGA 642
Db 140 ProLeuLeuLeuGlyIleCysPheIleGlyLeuAsnSerGlySerPhe-----Asn 156
Qy 643 GTTGTGTAATGTTATTTTCCCTGGTATCTGTCACCTATGAGGCGCCACCTTTGCAATATG 702
Db 157 ValIleLeuTrpAspLysGlyAspMetCysAspSerTyrGlyAlaPheLeuIleIle 176
Qy 703 TTG----- 705
Db 177 ValMetPheLeuThrIlyIleProValPheProPheHisGlyTrpLeuProLeuValHis 196
Qy 706 -----GCTATCCAGATCAAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 750
Db 197 AlaGluAlaSerSerProValSerIleIleLeuSerGlyTyrIleMetLysLeuGlyLeu 216
Qy 751 CTAGAAGTCTCTGCATCACCATCTGGCGGACATTTGGAGATCAGCTTCGCGCTCTGATT 810
Db 217 ValGlyLeuValArgLeuCysGlyTrpLeuLeuIleAsp-----Tyr 230
Qy 811 CTGTGCTCTCTCAGCCACTTTGAAATTGAAGCTGTGCGCTTCTAGTGTCAACTTC 870
Db 231 IleTyrTyrPheSerThrPheLeuLeuCysTyrSerValValTyrLeuVal----- 247
Qy 871 CTGATCATCTCTTGGAGCCCTGGATTAAGTTCTGGAGAAGTGTGGCCAGATGCGCAAT 930
Db 248 ---AlaAlaValPheGluCysAspSerIlyArgTrpLeuAlaTyrLeuSerLeuSerHis 266
Qy 931 AACATTGAGAAAACTTCAGCGGGTGGCAGCTCTGGTGGTGGCTG----- 975
Db 267 IleLeu-----IleGlyValCysIleLeuLeuThrSerThrTyrCys 280
Qy 976 -----ATTTCAGTCCACCTCTCTAT 996
Db 281 GlyAspTyrLeuAlaPheIleTyrCysLeuGlyHisGlyLeuSerValAlaLeuPhe 300
Qy 997 GCTGGCATCAACTCTCTGCTGGTGCAGCTTGGGAGGCTGCATCTGCT 1056

```

```

Db 301 MetValile-----TrpPheGlyTyrGlu----- 308
QY 1057 GTCGACAAAGGGCAGAACTGGGACATATGGCCCTGACCTATAGTGTGAGGTGGTAGAG 1116
Db 309 IleSerGlySerArgAnthrGlyIleLeuValLysIlePheGlyGlyLeuIleMet 328
QY 1117 AATGTGATCATGTCTTGGTTT-----AAGTTC 1146
Db 329 HisPheIleMetGlyPheValPheLeuAsnValCysGlyPheProProAlaLeuGlnPhe 348
QY 1147 TTGGA---GTGAAGTGTACTGAATTAC-----TGTCAATCTTGTGATGGCTTCAG 1197
Db 349 PheGlyGluLeuTrpLeuValIleAsnTyrIleThrLeuGlyAspIleIleSerLeuLeu 368
QY 1198 CTCATTATTGCTTATCTGATTTC-----ATTGGCTTCATGCTCTCTTCTTCCAG 1248
Db 369 LeuValSerIleTyrIlePheSerGlySerIleIleGlyPheIleIleTyrGlyLeuVal 388
QY 1249 TACTTCATCCATTCGGCTCCTCCTCCACCAATAA-----GTAGTAGACTACCTCCAT 1302
Db 389 IleCysSerProIleAnthrSerTyrGluTyrSerGlyLeuAspAsnPheLeuPhe 408
QY 1303 TGTGTC 1308
Db 409 CysIle 410

RESULT 15
Q94UX6 PRELIMINARY; PRT; 424 AA.
AC Q94UX6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN NAD4.
OS Schistosoma japonicum (Blood fluke).
OC Mitochondrion.
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6182;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20349913; PubMed=10889225;
RA Le T.H., Blair D., Agatsuma T., Humair P.P., Campbell N.J.,
RA Iwagami M., Littlewood D.T., Peacock B., Johnston D.A., Bartley J.,
RA Rollinson D., Herniou E.A., Zarlinga D.S., McManus D.P.;
RT "Phylogenies inferred from mitochondrial gene orders-a cautionary tale
RT from the parasitic flatworms.";
RL Mol. Biol. Evol. 17:1123-1125(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Le T.H., Blair D., McManus D.P.;
RT "Revisiting limited genetic variation within Schistosoma japonicum
RT populations.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -/- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
DR EMBL; AF412209; AAL12139.1; -.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ SEQUENCE 424 AA; 47727 MW; E65892281546EB93 CRC64;

Alignment Scores:
Pred. No.: 0.021 Length: 424
Score: 116.50 Matches: 90
Percent Similarity: 36.36% Conservative: 78
Best Local Similarity: 19.48% Mismatches: 155
Query Match: 4.58% Indels: 139
DB: 8 Gaps: 20

```

US-09-768-781-2 (1-1389) x Q94UX6 (1-424)

```

QY 133 ACTTTTCATTAGCATCCCTTTTCCACCTTTTGTACTGTGGGAGGCTGCATGCT 192
Db 18 ThrPheIlePheLeuIlePheTyrSerSerIleTrpValSerAspSerSerMetVal 37
QY 193 TTGTACATGTTAGATCTATCGAAGAATAGTGAACACTTACTGTGATGACATCACCTTT 252
Db 38 MetValGlyValIlystYrTyrLeuCysAspGlyLeuValIleIleAspThr----- 54
QY 253 TCITTTCTTTATGTTTTCATCCATTATGGTCCAGTTGACC---CTCATTTTGTCCACAGA 309
Db 55 -----LeuSerCysLeuMetIlePheLeuThrSerIleIleTrpLeuValLeu 70
QY 310 GATCTAGCCAAAGATAAACCGCTATCATTTATTTATGATCTAATCTCTTGGGACCTGTT 369
Db 71 TrpLeuValGlySerIlyAspIleValLeuPhe----- 81
QY 370 ATCAGATGTTTGGAGGCCATGATTAAGTAC-----CTCACACTGTGGAG 414
Db 82 IleSerValPheSerAlaMetLeuThrTyrValValSerAsnSerLeuValPheTrpPhe 101
QY 415 AAAGAGGAGCAGGAGGAGCCCTATGTCAGCTCACCCGAAAGAGATGCTAATAGATGCG 474
Db 102 PheTyrGluLeuSer-----IleIleSerAlaLeuTyrMetLeuIleValGly 117
QY 475 -----GAGGAGGTGCTAGATAATGGAGGTGGGCCACTCCATCCGAGCCCTG 522
Db 118 SerProTyrProGluArgTyrIleSerSerTyrTrpPheGlyGlyTyrIle----- 134
QY 523 GCTATGCACCGCAATGCCTACAAACGCTATGTCAGATCCAAAGCCTTCTCTGGGCTCAGTG 582
Db 135 -----LeuLeuSerSerVal 139
QY 583 CCCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCTGAGAGGTTCCTCCGGGTAGA 642
Db 140 ProLeuLeuLeuGlyIleCysPheIleGlyLeuAsnSerGlySerPhe-----Asn 156
QY 643 GTTGTGCTAATGTTATTTTCCCTGGTATCTGTCACTATGGGGCCACCTTTGCAATATG 702
Db 157 ValIleLeuTrpAspIlyGlyAspMetCysAspSerTyrGlyAlaPheLeuIleIle 176
QY 703 TTG----- 705
Db 177 ValMetPheLeuThrLysIleProValPheProPheHisGlyTrpLeuProLeuValHis 196
QY 706 -----GCTATCCAGATCAAGTAGATGACTACAGATTCGCTTGGGCCA 750
Db 197 AlaGluAlaSerSerProValSerIleIleLeuSerGlyTyrIleMetLysLeuGlyLeu 216
QY 751 CTAGAAGTCTCTGTCATCACCATCTGCGGACATGAGATCACTTCCCGCTCCTGATT 810
Db 217 ValGlyLeuValArgLeuCysGlyTrpLeuLeuIleAsp-----Tyr 230
QY 811 CTGGTGTCTTCTCAGCCACTTTTGAATTTGAAGGTGTGCTTCTCTAGTGTCTCAACTTC 870
Db 231 IleTyrTyrPheSerThrPheLeuLeuLeuCysTyrSerValValTyrLeuVal----- 247
QY 871 CTGATCATCTCTTTGAGCCCTGGATTAAGTCTCTGAGAAGTGTGCGCCAGATGCCCAAT 930
Db 248 ---AlaAlaValPheGluCysAspSerIlyArgTrpLeuAlaTyrLeuSerLeuSerHis 266
QY 931 AACATTGAGAAAAACTTCAGCCGGGTGCGCACTCTGGTGTCTCTG----- 975
Db 267 IleLeu-----IleGlyValCysIleLeuLeuThrSerThrTyrCys 280
QY 976 -----ATTTCAGTCACCATCTCTAT 996
Db 281 GlyAspTyrLeuAlaPheIleTyrCysLeuGlyHisGlyLeuSerValAlaLeuLeuPhe 300
QY 997 GCTGGCATCAACTCTCTTGTCTGCTGAGCTTTGAGTTGAGTTGGCAGACAGATCTC 1056
Db 301 MetValIle-----TrpPheGlyTyrGlu----- 308
QY 1057 GTCGACAAAGGGCAGAACTGGGACATATGGGCTGACCTACTAGTGTGAGGTGGTAGAG 1116

```

```
Db 309 ILeSerGlySerArgAsnTrpGlyIleLeuValLysIlePheGlyGlyGlyLeuIleMet 328
QY 1117 AATGTGATCATGGTCTTGGTWTTC-----AGTTTC 1146
Db 329 HisPheIleMetGlyPheValPheLeuAsnValCysGlyPheProProAlaLeuGlnPhe 348
QY 1147 TTTGGA--GTGAAAGTGTACTGAATTAC-----TGTCAATTCCTTGATTCGCTTCAG 1197
Db 349 PheGlyGluLeuTrpLeuValIleAsnTyrIleThrLeuGlyAspIleIleSerLeuLeu 368
QY 1198 CTCATTATTGCTTATCTGATTTC-----ATTGGCTTCATGCTCCTTTCTTCCAG 1248
Db 369 LeuValSerIleTyrIlePheSerGlySerIleIleGlyPheIleIleTyrGlyLeuVal 388
QY 1249 TACTTGCAATTCGCGTCACTCTTACCCATAAT-----GTAGTAGACTACCTCCAT 1302
Db 389 IleCysSerProIleAsnThrSerTyrGluTyrSerGlyGlyLeuAspAsnPheLeuPhe 408
QY 1303 TGTGTC 1308
Db 409 CysIle 410
```

Search completed: April 1, 2003, 08:57:34
Job time : 108.5 secs